

基于一种改进神经网络的数据手套手势识别

王新宇, 孙冠, 韩冬, 张婷

北京理工大学 自动化学院, 北京 100081

E-mail: xsunguanx@gmail.com

摘要: 本文首先介绍了用于进行手势识别的数据手套, 并针对基于数据手套的手势识别提出了一种改进的神经网络算法。利用神经网络方法实现了聚类分析, 改进后的聚类神经网络算法既具有聚类分析直观、易行的优点, 又具有神经网络善于学习、适应性强的优点。实验证明, 此算法显著提高了手势识别的智能化水平。

关键词: 聚类分析、神经网络、数据手套、手势识别

Data Glove Gesture Recognition Based on an Improved Neural Network

Wang Xinyu, Sun Guan, Han Dong, Zhang Ting

School of Automation, Beijing Institute of Technology, Beijing 100081

E-mail: xsunguanx@gmail.com

Abstract: This paper describes the data glove used for gesture recognition, and proposes an improved neural network algorithm for gesture recognition based on data glove. The neural network is used to achieve clustering, which is not only intuitive as the cluster analysis, but also has the ability to learn from the data as the neural network. Experimental results show that this algorithm significantly improved the intelligence level of gesture recognition.

Key Words: Cluster Analysis, Neural Network, Data Glove, Gesture Recognition

1 引言(Introduction)

当今科技的飞速发展离不开计算机的广泛应用, 计算机已经渗透于生产和生活的各个方面, 实现更好的人机交互方式也日益成为科研领域的热点。手势, 作为人类最常用的表达方式, 是一种自然、直观、高效、易于学习的人机交互手段。

目前的手势识别技术分为基于数据手套和基于计算机视觉两大类。其中基于数据手套进行手势识别的优点是输入数据量小、识别速度快、识别准确率高、不易受外界条件影响。本文涉及的研究以数据手套为基础, 它可以实时地测量人手各个关节角度的变化和手部的运动状态, 并将这些信息送往计算机, 由计算机进行解算和理解, 达到识别手势的目的。

2 数据手套简介(Introduction to Data Glove)

研究中采用的是作者自主研发的数据手套。该手套上安装有位于手指部分的五个应变片弯曲传感器和位于手背部分的三轴加速度计和三轴磁阻, 它们分别完成对五个手指的弯曲角度的测量以及对手部运动状态的检测。弯曲传感器安装方式见图1。

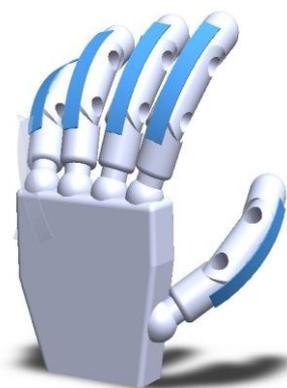


图 1: 弯曲传感器安装示意图

位于手套手背部位的单片机系统接收到各个传感器检测到的数据, 进行相应的AD转换和处理后发送至上位机, 由上位机根据所得传感器信息进行解算, 最终得到用户手的姿态和动作信息。数据手套的电路结构图见图2。

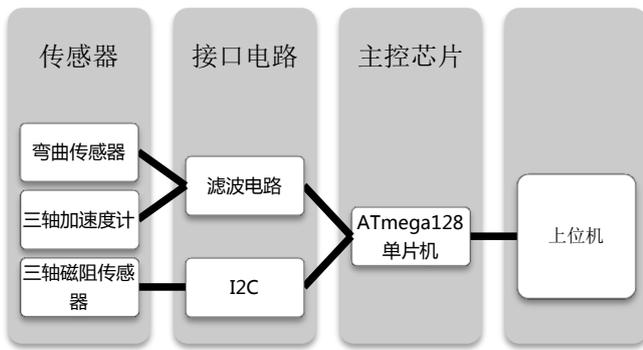


图 2: 数据手套电路结构图

对于静态手势的识别，需要的信息是五个手指的弯曲程度及三轴加速度计的测量值，上位机通过对这五个数据的解算完成手势识别，常用的算法有阈值匹配判断、BP神经网络等，这些算法简单易行，但由于数据手套的结构限制以及弯曲传感器的特性等问题，采集到的数据往往与手指的实际弯曲情况有偏差，并且线性度不好；同时，数据还会因使用者的手型、动作习惯等而有差异。因此，上述算法往往在准确率、适应性等方面上有所不足。针对这种情况，本文将聚类分析和神经网络相结合，作为手势识别的算法。

3 一种改进的神经网络(An Improved Neural Network)

3.1 基本原理 (Basic Principles)

本文把聚类分析与神经网络相结合，以聚类分析指导神经网络的训练和识别。使改进后的神经网络既具有聚类分析直观、易行的特点，又具有神经网络学习能力强、适应性强的特点。

以神经网络的方式来实现聚类分析，使神经网络不再是黑箱结构，而具有实际直观的物理意义；同时利用神经网络的学习能力，可动态调整各输入权值，优化分类结果，最终达到比较理想的手势识别效果。

3.2 结构设计 (Structure Design)

使用数据手套对手势进行识别的过程是一个将输入的手势进行分类的过程，本文分类采用的基本方法是以多维空间中的加权欧几里得距离为判据的聚类分析。以神经网络的方式实现此聚类分析，需在网络中实现距离分量计算、各维度权值确定以及判据距离计算的功能。

根据以上分析，设计如下图所示的神经网络用于手势的分类识别。该网络在结构上与自组织特征映射神经网络相似，分为输入层和输出层，如图3所示。

输入层：该层采集传感器数据并进行运算处理后送入输出层。由于在进行手势识别时需要各手指的弯曲程度和手部姿态的信息，即需要处理手套中五个弯曲传感器和一个三轴加速度计传回的八路数据，所以该层共设置八个节点。输入的八个量可以看做八维空间中的一个特征向量。

连接：此神经网络采用全连接的方式，即每一个输出节点都与输入层的八个输入节点连接，每个连接都有对应的输入权值 ω 。每个输出节点对应的八个 ω 表征了八个输入量对于此输出节点的不同重要程度。输入权值在训练的过程中进行调整，从而使识别过程中关注的重点集中在能够表征一种手势的某一个或几个量上，实际上这也是人类判断手势时的一个基本特点。

输出层：该层也可称为竞争层，其中的每个输出节点根据各输入数值和各输入权值运算后输出。输出节点数量与预期识别的手势数量相同，即每一个输出节点对应一种手势。

由以上的结构描述可知，该网络较之普通的神经网络，结构简单、运算量小，并且每一层有明确的物理意义。

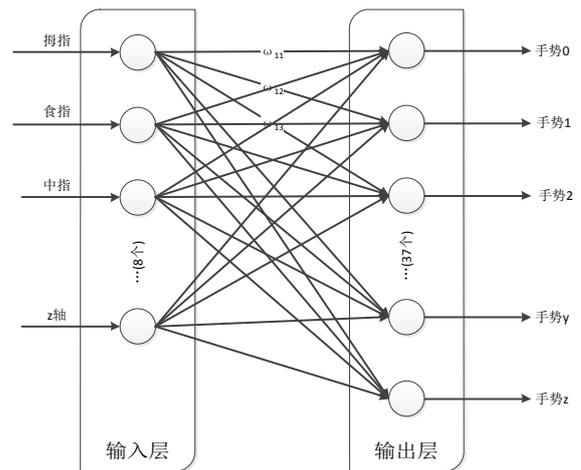


图 3: 神经网络结构图

3.3 算法设计 (Algorithm Design)

结合前文所述的神经网络结构设计手势识别算法，包括训练算法和识别算法，在进行手势识别前先通过训练算法对神经网络进行训练，调整各输入权值，训练完成后即可输入手势进行识别。具体算法流程图如图4所示：

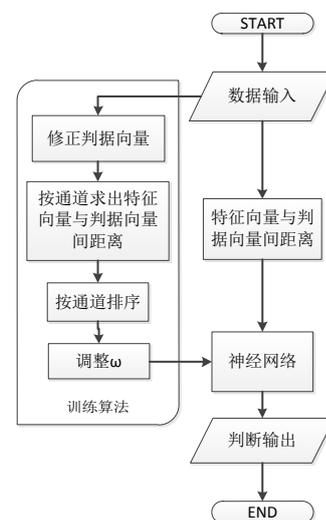


图 4: 算法流程图

3.3.1. 训练算法 (Training Algorithm)

网络中需要学习的参数主要是各通道的输入权值 ω 。调整权值 ω 的意义在于使识别过程中关注的重点集中在能够表征一种手势的某一个或几个量上,实际上这也是人类判断手势时的一个基本特点,例如在识别手势“一”时,对食指的关注程度要明显高于其他手指,也高于对手部姿态的关注程度。

以手语字母A的手势学习为例,首先记录一个手势A的特征向量作为初始识别判据向量,在之后的训练过程中,再次做出手势A,将新的特征向量记录下来,并与当前手势A的判据向量进行比较,按下式求出新的判据向量:

$$J_{new} = \frac{J_{old} + S}{2} \quad (1)$$

其中:

S 为新的输入特征向量;

J_{old} 为之前的相应判据向量;

J_{new} 为更新后的判据向量。

根据新得到的判据向量和当前输入的特征向量之差的平方在各通道上的分量,调整每个通道上的输入权值 ω_i 。权值修改的步长取决于以下几个数值:

- 某一通道与该向量均值的差的平方的大小;
- 随着训练次数指数衰减的系数;
- 开始训练时选定的步长。

具体调整方法如下:

$$\omega'_i = \omega_i - l \left[(s_i - j_i)^2 - \sum_{n=0}^7 \frac{(s_n - j_n)^2}{N} \right] e^{-\frac{t}{\tau}} \quad (2)$$

其中:

ω'_i 为新的权值向量的第*i*通道;

ω_i 为过去的权值向量的第*i*通道;

s_i 为新的输入特征向量 S 的第*i*通道;

j_i 为判据向量 J_{new} 的第*i*通道;

l 为训练步长;

t 为训练次数;

τ 为收敛速度。

对于权值有范围有一下限制:

- 所有通道权值之和需要归一化;
- 所有权值都必须是非负的。

将上文描述的学习过程重复若干次,可得到较为理想的权值向量 ω 。对每一个要识别的手势,都进行以上的训练过程,得到每一个手势所对应的适当的八个输入权值。

3.3.2. 识别算法 (Recognition Algorithm)

使用上文中所描述的聚类神经网络进行手势识别时,将数据手套五路弯曲传感器的值和三轴加速度计的值分别从各输入节点输入。

在输入层,由传感器采集回的八路数据,得到目前的输入八维特征向量,计算其与训练后的判据向量的距离在各通道上的分量,得到各通道传递给输出层的数据,具体计算公式为:

$$a_i = (s_i - j_i)^2 \quad (3)$$

其中:

a_i 为神经网络输入向量的第*i*个分量

在输出层,使用学习后得到的各通道权值,以及各通道的输入量,使用以下公式计算每个输出节点的输出值:

$$a' = \min(\omega_i a_i) \quad (4)$$

其中:

a' 为最终获胜的输出节点的输出值。

输出值 a' 实际上也可以理解为八维空间中输入的特征向量到判据向量的加权欧几里得距离,以此距离作为聚类分析的判据,距离越小,说明输入的特征向量与代表某一手势的判据向量越接近,因此输出值最小的节点获胜,识别结果为此输出节点所对应的手势。

4 结论与分析(Conclusion and Analysis)

实际的应用系统中,在Windows 7操作系统环境下,采用Visual C++语言编程,实现了如上所述神经网络,配合数据手套进行训练后,可以实现对数字0至10,以及字母A至Z的手势的正确识别。图5中所示为根据手势识别出不同的数字和字母。

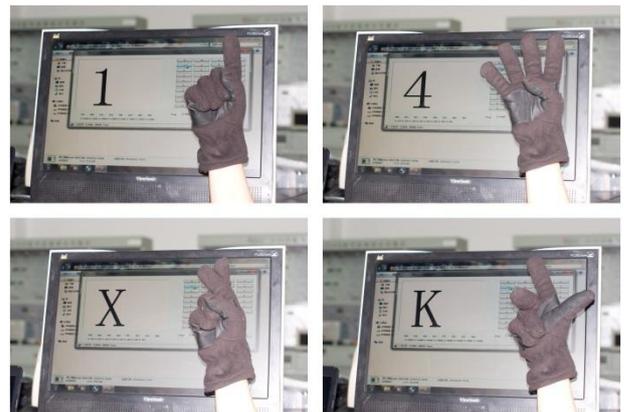


图 5: 识别效果图

如图6、7、8为对数字“一”、数字“四”和字母“X”的手势进行训练时,其各自对应的八个输入权值变化过程。可见,随着训练次数的增加,原本相等的八个输入权值大小发生改变并逐渐稳定下来,形成新的大小排布,以使识别过程中关注于能够表征此手势的某几个输入量上,增加了识别的准确性和宽容性。

图6为对数字“一”的手势对应的八个输入权值的变化过程,代表食指通道的输入权值增大,其他通道减小,与人类在识别手势“一”时更为关注食指的习惯相符。

图7为对数字“四”的手势对应的八个输入权值的变化过程，代表食指、中指、无名指和小指的输入权值增大，其他通道减小，与人类在识别手势“四”时更为关注此四根手指的习惯相符。

图8为对字母“X”的手势对应的八个输入权值的变化过程，代表食指和中指通道的输入权值增大，其他通道减小，与人类在识别手势“X”时更为关注食指和中指的习惯相符。

图9为对字母“K”的手势对应的八个输入权值的变化过程，代表拇指、食指和中指通道的输入权值增大，其他通道减小，与人类在识别手势“K”时更为关注拇指、食指和中指的习惯相符。

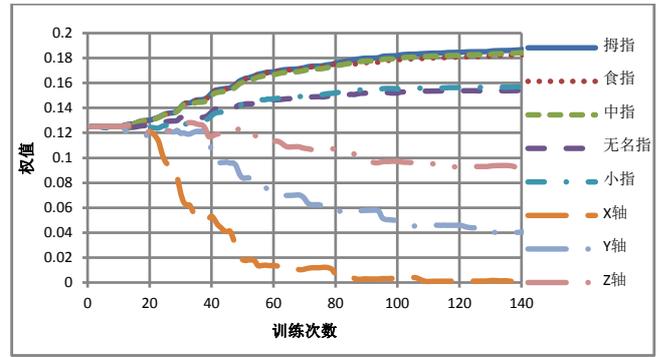


图 9: 手势“K”

本文所涉及的研究将聚类分析与神经网络相结合，具体来说就是以神经网络的方式实现了以距离为判据的聚类，得到的聚类神经网络结合了聚类分析直观易行和神经网络善于从知识中学习的优点。将此聚类神经网络应用于基于数据手套的手势识别，其特点为：输入数据量小、识别速度快、训练简单、训练速度快、可适应性好、识别正确率高，显著提高了手势识别的智能化水平。

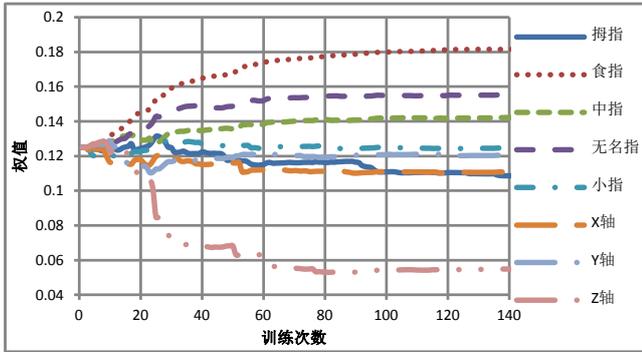


图 6: 手势“一”

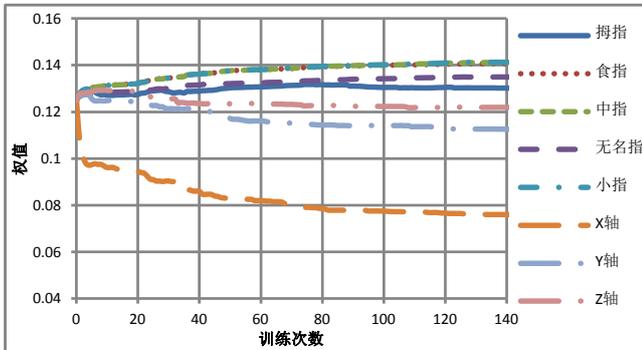


图 7: 手势“四”

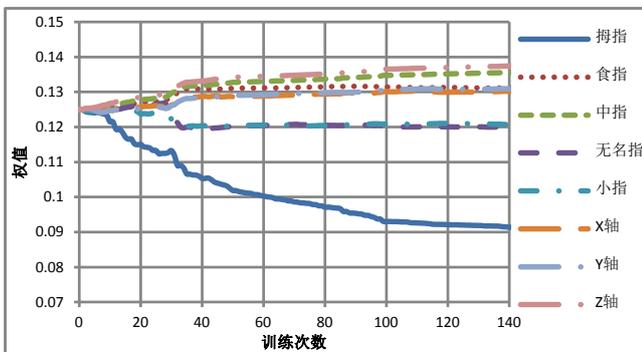


图 8: 手势“X”

参考文献(References)

- [1] Fifth Dimension Technology Corporation. 5DT data gloves serial manual [R]. [s. l.]: Fifth Dimension Technology Corporation, 2004.
- [2] 施教芳, 陈珏, 赵鹤鸣. 基于节点对BP网络的数据手套手势识别[J]. 系统仿真学报. 2008(6), Vol.4, No.3.
- [3] Simon Haykin. 神经网络原理[M]. 北京: 机械工业出版社, 2006
- [4] 飞思科技产品研发中心. 神经网络理论与Matlab7实现[M]. 北京: 电子工业出版社, 2005.
- [5] 李弼程, 邵美珍, 黄洁. 模式识别原理与应用[M]. 西安: 西安电子科技大学出版社, 2008